

# Prediction accuracy and sloppiness of log-periodic functions

DAVID S. BRÉE\*†¶, DAMIEN CHALLET‡¶ and PIER PAOLO PEIRANO§¶

†Institute for Social Studies, University of Warsaw, Warsaw, Poland

‡Physics Department, University of Fribourg, Fribourg, Switzerland

§Capital Fund Management, Paris, France

We show that log-periodic power-law (LPPL) functions are intrinsically very hard to fit to time series. This comes from their sloppiness, the squared residuals depending very much on some combinations of parameters and very little on other ones. The time of singularity that is supposed to give an estimate of the day of the crash belongs to the latter category. We discuss in detail why and how the fitting procedure must take into account the sloppy nature of this kind of model. We then test the reliability of LPPLs on synthetic AR(1) data replicating the Hang Seng 1987 crash and show that even this case is borderline regarding the predictability of the divergence time. We finally argue that current methods used to estimate a probabilistic time window for the divergence time are likely to be over-optimistic.

**Keywords:** Financial time series; Mathematical finance; Modelling asset price dynamics; Power laws

**JEL Classification:** C2, C5, C22, C52

## 1. Introduction

Log-periodic functions have received much attention because of the claim that they could be used to predict the times of singularities. While they are known to occur in hierarchical discrete scale-free networks (Sornette 1998, Suchecki and Holyst 2005), they have been claimed to have been observed in many types of natural time/size series: earthquakes (Newman *et al.* 1995, Sornette and Sammis 1995, Borodich 1997, Tanaka 2002, Kawada *et al.* 2008), icequakes (Pralong *et al.* 2005, Faillettaz *et al.* 2009), forest fires (Malamud *et al.* 2005) as well as evolutionary trees (Nottale *et al.* 2000), although such claims have not gone unchallenged (Huang *et al.* 2000). But the most noticed application of such functions is to speculative bubbles of stock indices (Sornette *et al.* 1996, Sornette and Johansen 1997, Johansen *et al.* 2000, Sornette and Johansen 2001, Bastiaansen *et al.* 2009), foreign exchange rates

(Matsushita *et al.* 2005, Johansen and Sornette 2010), real estate (Zhou and Sornette 2006, 2008) and commodity prices (Drożdż *et al.* 2008a,b, Sornette *et al.* 2009) as well as downward spirals during the burst of the bubble (Johansen and Sornette 2001, Lillo and Mantegna 2004). Given the importance of such phenomena, and the possibly important consequences of finding a universal model that could be applied to this remarkable variety of bubbles, it is of course necessary to assess the statistical significance of Log-Periodic Power Laws (LPPLs) regarding crashes, i.e. the predictive power of log-periodic functions in this context. The question is still unsettled as of yet (Laloux *et al.* 1999, Feigenbaum 2001, van Bothmer 2003, Chang and Feigenbaum 2006, Sornette and Zhou 2006, Gazola *et al.* 2008). Problems are indeed numerous: what definition of a bubble and a crash to adopt (Jacobsen 2009, Lin *et al.* 2009),<sup>⊥</sup> should the price in a bubble always be increasing (Bothmer and Meister 2003), should one

\*Corresponding author. Email: dbree@cs.man.ac.uk

¶Previously at ISI Foundation, Turin, Italy.

<sup>⊥</sup>For example, Lin *et al.* (2009) recently claim that the LPPL model has few false positives on the grounds that all five of the bubbles that they identified on the S&P500 between 3 January 1950 and 21 November 2008 “preceded well-known crashes.” Using a criterion for a crash of a drop by >15% within 12 weeks identifies 12 crashes in this period, but only two of these (25/8/1987, 17/7/1998) were among the five claimed by Lin *et al.* (2009, table 3); the other three “well-known crashes” (in 1994, 1997 and 1999) do not meet these quite lenient crash criteria.

Table 1. Eigenvalues,  $\lambda$ , and associated eigenvectors of the best fit of real prices for the 1987 crash. Components of absolute value larger than 0.1 are in bold face.

$\lambda$	$A$	$B$	$C$	$t_c$	$\alpha$	$\omega$	$\phi$
$1.11 \cdot 10^9$	-0.00004	-0.00022	0.03198	0.00012	<b>0.99943</b>	0.01085	0.00158
$1.21 \cdot 10^7$	0.00001	0.00005	<b>0.99949</b>	-0.00016	-0.03200	0.00176	0.00060
$1.49 \cdot 10^6$	-0.00009	-0.00018	-0.00219	-0.00146	-0.01090	<b>-0.98737</b>	<b>-0.15804</b>
$2.89 \cdot 10^2$	0.00862	0.01599	0.00025	<b>-0.27339</b>	-0.00012	<b>0.15239</b>	<b>-0.94958</b>
$3.13 \cdot 10^1$	0.03973	0.08183	-0.00024	<b>-0.95699</b>	0.00017	-0.04192	<b>0.27063</b>
$2.48 \cdot 10^{-1}$	<b>0.39153</b>	<b>0.91509</b>	-0.00003	0.09615	0.00020	0.00102	-0.00855
$6.50 \cdot 10^{-3}$	<b>-0.91915</b>	<b>0.39367</b>	-0.00001	-0.01353	0.00005	-0.00031	0.00213

impose constraints on the fitted parameters (Johansen and Sornette 2001), where to start a fit of a bubble,<sup>†</sup> what test of goodness of fit to use,<sup>‡</sup> why having different lengths of the data window greatly affects the parameters of the best fit of the LPPL to the data (Feigenbaum 2001), why leaving out a few data points can alter the parameters of the best fit sufficiently to change a no/bubble decision (Lin *et al.* 2009), and why is the fitting error very sensitive to small (but not large!) changes in one of the parameters of the model (Brée and Joseph 2010)?

What contributes to most if not all of these difficulties is that a stable best fit of an LPPL to the data is very hard to determine. Here we aim to show that this comes from the fact that LPPLs belong to the family of sloppy functions, a terminology introduced in a series of papers by Sethna *et al.* (Brown and Sethna 2003, Frederiksen *et al.* 2004, Water *et al.* 2006, Gutenkunst *et al.* 2007); we will discuss in detail what this means when applying LPPLs to noisy time series.

## 2. Sloppiness

Let us denote by  $p(t)$  the time series to be fitted,  $f$  the fitting function and  $\Phi$  the set of parameters. Least-squares fits minimise  $S = \sum_{t=t_0}^{t_1} [f_{\Phi}(t) - p(t)]^2 / (t_1 - t_0 - n)$ , where  $t_1 < t_c$ , the time when the singularity (crash) occurs, and  $n$  is the number of free parameters. The best fit  $\hat{\Phi}$  to some given data corresponds by definition to the minimum of  $S$ , therefore close to  $\hat{\Phi}$ ,  $S \simeq S_0 + \sum_{x,y \in P} (\partial^2 S / \partial x \partial y)|_{\Phi=\hat{\Phi}} (x - \hat{x})(y - \hat{y})$ . Assuming that  $\hat{\Phi}$  does not sit on a boundary of the parameter space, the curvature of  $S$  in a neighbourhood of  $\hat{\Phi}$  is positive; as a consequence, the Hessian  $(\partial^2 S / \partial x \partial y)|_{\Phi=\hat{\Phi}}$  is positive-definite, thus all its eigenvalues are positive.

As shown recently in a series of papers (e.g., Brown and Sethna (2003), Frederiksen *et al.* (2004), Water *et al.* (2006) and (Gutenkunst *et al.* 2007)), sloppy models are characterised by a separation of Hessian eigenvalues by orders of magnitude, which is all the more likely and evident when the models have many parameters. Combinations of parameters corresponding to larger

eigenvalues are called stiff, while those corresponding to small eigenvalues are called sloppy. In other words, varying slightly a stiff parameter combination has a large influence on  $S$ , while changing combinations of sloppy parameters does not modify  $S$  substantially. This has two consequences, discussed in detail in the following sections: first, fitting sloppy functions must be done carefully; second, out-of-sample predictions from the best-fit values of sloppy parameters may be imprecise: as the noise from sample to sample changes, the fitted values of sloppy parameters are likely to change greatly.

Let us apply this reasoning to the fitting of financial index prices  $p(t)$  with log-periodic functions, as used originally in this context by Sornette *et al.* (1996),

$$f_{LP}(t) = A + B(t_c - t)^{\alpha} [1 + C \cos(\omega \log(t_c - t) + \phi)].$$

This is a seven-parameter fit, but as already noted in the original paper, minimising  $S$  with respect to  $A$ ,  $B$ , and  $C$  yields linear equations, which reduces the nonlinear part of the fitting problem to four parameters. However, sloppiness concerns *a priori* all seven parameters. This is why we shall keep them all, i.e.  $\hat{\Phi} = \{\hat{A}, \hat{B}, \hat{t}_c, \hat{\alpha}, \hat{C}, \hat{\omega}, \hat{\phi}\}$ , in order to give a fuller account of sloppiness. Once we understand what respective importance  $A$ ,  $B$ , and  $C$  have in  $S$ , we will be able to focus on the other parameters.

It turns out that log-periodic functions are very sloppy: every crash we fitted resulted in a clear separation of eigenvalues by orders of magnitude. Let us take, for example, the 1987 crash in the Hang Seng index. The eigenvalues and eigenvectors of the Hessian of the best fit obtained by using the Levenberg–Marquardt algorithm to fit the 834 days preceding the crash, and retaining the best of a set of 20,000 initial conditions, are shown for the best fit to real prices in table 1 and to log prices in table 2.

These tables contain several relevant pieces of information. First, the largest eigenvalue is at least nine orders of magnitude larger than the smallest one, a definite signature of sloppiness; in addition, the eigenvalues are well spread over these orders of magnitudes. The associated eigenvectors confirm the wisdom that the stiffer a direction, the more likely that it is close to an axis, and

<sup>†</sup>While the beginning of a bubble is supposed to be given by the lowest value since the previous crash, it may be moved from this minimum to some later date; in Johansen and Sornette (2001) the beginnings of half of the eight bubbles on the Hang Seng up to 1998 were thus moved (Brée and Joseph 2010).

<sup>‡</sup>Lomb periodograms (van Bothmer 2003) or the residuals (Gazola *et al.* 2008, Lin *et al.* 2009) have been proposed, with mixed results.

Table 2. Eigenvalues,  $\lambda$ , and associated eigenvectors of the best fit of the log price for the 1987 crash. Components of absolute value larger than 0.1 are in bold face.

$\lambda$	$A$	$B$	$C$	$t_c$	$\alpha$	$\omega$	$\phi$
$4.16 \cdot 10^4$	0.00621	<b>0.99935</b>	-0.00141	-0.00001	-0.03546	-0.00241	-0.00035
$1.12 \cdot 10^0$	<b>0.38694</b>	0.01331	0.01048	0.00036	<b>0.38499</b>	<b>0.82776</b>	<b>0.12877</b>
$5.12 \cdot 10^{-1}$	0.07433	0.00942	<b>0.95117</b>	-0.00080	<b>0.25177</b>	<b>-0.16023</b>	-0.02445
$3.56 \cdot 10^{-1}$	<b>-0.73357</b>	-0.00726	<b>0.24417</b>	0.00276	<b>-0.37764</b>	<b>0.50250</b>	0.08409
$1.03 \cdot 10^{-1}$	<b>-0.55369</b>	0.03139	<b>-0.18851</b>	0.00156	<b>0.80280</b>	<b>-0.10934</b>	-0.02146
$4.48 \cdot 10^{-5}$	-0.00226	-0.00015	0.00260	<b>-0.20087</b>	-0.00560	<b>0.15421</b>	<b>-0.96738</b>
$1.86 \cdot 10^{-5}$	0.00198	-0.00002	0.00078	<b>0.97961</b>	-0.00067	0.02986	<b>-0.19864</b>

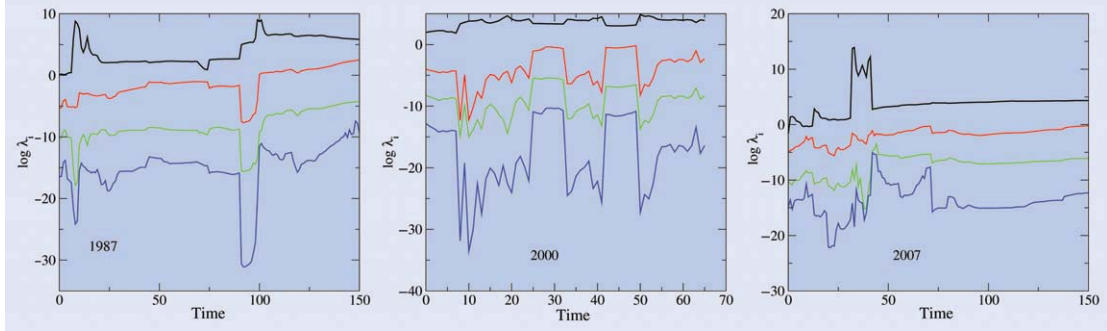


Figure 1. Eigenvalues associated with the four parameters requiring nonlinear fitting as a function of time in days immediately before  $t_c$  for the 1987, 2000, and 2007 crashes (log prices).

reversely for sloppy eigenvalues (Gutenkunst *et al.* 2007). Next, the eigenvectors vary from crash to crash and can be quite different between real and log-prices: for the 1987 crash, the linear-fit parameters ( $A$ ,  $B$ , and  $C$ ) are completely disconnected from the other ones only in the case of real prices; curiously, this is not systematic, as both log and real prices of the 1997 crash lead to disconnected eigenvectors, for instance. When the eigenvectors associated with  $A$ ,  $B$ , and  $C$  are not completely disconnected from the other four parameters, one should not fit them separately when estimating the error associated with  $t_c$ , for instance (see the Discussion section below and Brown and Sethna (2003) and Frederiksen *et al.* (2004)).

We are, of course, chiefly interested in the role of  $t_c$ : this crucial parameter turns out to be one of the most sloppy parameters and, as a consequence, its associated eigenvector is not along the  $t_c$  axis but also comprises the phase  $\phi$  and the frequency  $\omega$ , meaning that in order to fit  $t_c$  precisely, one should take this inter-dependency into account properly, which is not the case in the state-of-the-art papers on the topic that all rely on the Levenberg-Marquart algorithm (see below for remedies).

Sloppiness is intrinsic to the LPPL equation, not only to the 1987 crash, nor just to dangerous times just before a crash. In order to convince oneself of this important point, figure 1 plots the four eigenvalues associated with the parameters requiring a nonlinear fit as a function of time in the 150 days preceding the 1987, 2000, and 2007 crashes on the Hang Seng, chosen randomly; it is obvious that the eigenvalues are well-spaced and that their typical spacing stays very large in the whole time series; their

structure is also constant, with no crossing of eigenvalues. We have found the same behaviour for all the crashes investigated. Note, however, that some crashes lead to more sloppy fits than others, i.e. with an even larger eigenvalue separation.

It should be noted that, in principle, some sloppy models can be unslopped by a suitable change of fitting functions. For instance, fitting a function in  $[0,1]$  with a sum of exponentials is known to be ill-posed (van den Bos and Swarte 1993). However, using Hermite's polynomials lifts the sloppiness of exponentials (Water *et al.* 2006). Unfortunately, this approach relies on a symmetry assumption between the parameters that does not hold for LPPL.

Sloppiness has important consequences and, despite its negative connotation, these are not only negative. However, being aware that LPPLs are sloppy models helps understand several important aspects of making predictions with an LPPL, in particular with respect to the uncertainty associated with the most sloppy parameters; this will be discussed in the next few sections.

### 3. Consequences of sloppiness

#### 3.1. Sensitivity of $t_c$

The main result of the previous section is that not only are LPPL functions sloppy, but that varying  $t_c$  together with  $\phi$  has little influence on the squared residuals. Reversely, changing the input slightly will vary  $t_c$  tremendously. This explains first why the diagnostic of a bubble is sometimes

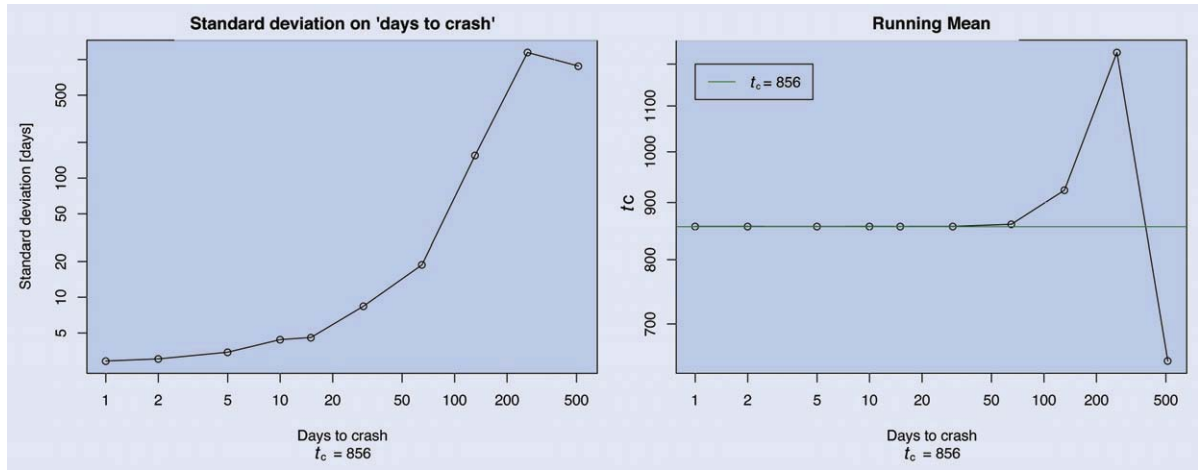


Figure 2. Average and standard deviation of the crash time estimate  $\hat{t}_c$  for synthetic data with AR(1) noise and parameters reproducing the 1987 Hang-Seng crash.

sensitive to the addition or deletion of a single data point. By extension, the sensitivity of  $t_c$  to noise must be investigated and one must understand how reliable the fits of LPPL to noisy data can be.

Quite tellingly, early papers using LPPL to predict various kinds of crashes used only a single fit, which, of course, is problematic in the light of sloppiness. Recent papers try to build a probabilistic window for  $t_c$  (Bastiaensen *et al.* 2009, Jiang *et al.* 2010, Yana *et al.* 2010). The problem one faces is to estimate a probability distribution for  $t_c$  from a *single* noisy time series. The methods consist essentially of varying the beginning and end of the time series, thereby obtaining a distribution of fitted values for  $t_c$ . But this only happens because LPPLs are sloppy and because  $t_c$  is one of the least relevant variables in the fit. Thus, this probabilistic method uses the intrinsic imprecision of LPPL regarding  $t_c$ . This is the positive side of sloppiness. The negative side is, of course, that the imprecision on  $t_c$  is *a priori* very large. In addition, there is no real guarantee that the distribution of  $t_c$  thus obtained corresponds to anything meaningful. As we shall explain below, special methods have been devised for sloppy functions that are able to give reliable probability distributions for fitted variables from a single time series.

### 3.2. Fitting LPPL

First, using simple fitting algorithms is bound to be problematic for sloppy functions (see, e.g., the discussion of Brown and Sethna (2003)), as most of them approximate to first order the cost function variation when trying to find the next move in the parameter space. In the case of sloppy functions, however, one needs to take into account not only the gradient, but also the curvature of the cost landscape by computing the eigenvectors and following them, which is computationally more costly. A computational compromise is the Levenberg–Marquart method (used by researchers studying LPPL ever since the original paper), which approximates the Hessian with a

product of gradients, thus implicitly assuming that the eigenvectors do not deviate much from the axes. While this is a reasonable approximation as regards some eigenvalues, as seen in tables 1 and 2, it breaks down, in particular for  $t_c$ : this means that reaching a correct estimate of  $t_c$  requires more sophisticated methods, such as the Rosenbrock method (Rosenbrock 1960) or the trust region algorithm (Byrd *et al.* 1987), at the cost of computational time. In this paper, we will restrict our attention to the performance and pitfalls of Levenberg–Marquart, hence applying such methods is beyond the scope of this paper.

### 3.3. Fitting full log-periodic functions with AR(1) noise

Of the recent progress, the residuals have been shown to be AR(1) (Gazola *et al.* 2008, Lin *et al.* 2009). It makes sense, therefore, to create artificial data with AR(1) noise. Let us consider the very simple case where one adds some noise to a pure log-periodic function and applies a fitting procedure. More specifically, we fit  $f_{LP}(t) + \sigma\eta(t)$ , where  $\eta$  follows an auto-regressive process  $\eta(t) = \eta(t-1)(1-\lambda) + \epsilon(t)$ , where  $\epsilon \sim \mathcal{N}(0, 1)$ ,  $\lambda$  is the memory loss and  $\sigma$  tunes the strength of the fluctuations. AR(1) noise that mimics the fit of LPPL functions to the 1987 crash is obtained with  $\lambda = 0.06$  and  $\sigma = 25$ . A natural test of the predicting power of the fit to  $f_{LP}$  is to consider a time series that starts at  $t=1$  and keeps expanding until  $t_c$ . We created 1000 such samples and computed the averages of the fitted parameters for increasing time series length. The average estimates of the parameters, quite remarkably including  $t_c$ , do converge to the true value at about 60 time steps (2.7 trading months) before the crash itself (figure 2). Thus it turns out that fitting an LPPL to synthetic data generated by an LPPL with a level of noise comparable to that of real markets is possible and that the average estimate of  $t_c$  behaves very well ahead of  $t_c$ . Therefore, one concludes that Levenberg–Marquart works well for estimating average parameter values for



synthetic data with many samples. Then a natural crash warning is obtained when the average of  $t_c$  stabilises.

However, when given a single run, predicting  $t_c$  is much more difficult: the standard deviation on  $t_c$  is about a half of  $t_c - t$ . Hence, since the residuals are Gaussian distributed (we have checked that this is the case), the 20–80% confidence window, as chosen in recent papers on predictions with LPPL (Bastiaensen *et al.* 2009, Jiang *et al.* 2010, Yana *et al.* 2010), corresponds to a width of about one standard deviation either side of the mean, hence ranges from  $(t + t_c)/2$  to  $(3t_c - t)/2$ , while the 5–95% confidence window (two standard deviations either side of the mean) ranges from  $t$  to  $2t_c - t$ . Therefore, when a crash warning is issued, the crash can occur any day at 95% confidence. Hence, predicting the date of a divergence is hard, even when the underlying time series is a real LPPL. The 1987 crash was chosen because LPPL fits it better than other crashes. The parameters associated with other crashes yield worse results.

#### 4. Discussion

Given the attention devoted to LPPL and despite recent technical developments, it is important to realise how sloppy this kind of function is. The sloppiness of LPPLs implies that special care must be taken when estimating the uncertainty on  $t_c$ . The leap of faith of LPPL regarding bubbles is not the log-periodic nature of oscillations, but to try to fit data with functions that contain a divergence. Thus the discussion on  $t_c$  is largely disconnected to the nature of the oscillations, as it is only related to a way to describe super-exponential growth. Obviously, one can fit real data with a function that does not contain oscillations by setting  $C=0$ , thus focusing on the super-exponential growth. We tried it on real data and while the precision on  $t_c$  is slightly worse than that obtained with a LPPL, it is a simpler method of obtaining an estimate for  $t_c$ . The fit of synthetic data with AR(1) noise is most revealing for several reasons: it first shows that the Levenberg–Marquart algorithm is adequate for noisy *synthetic* data, that is when the underlying function is of LPPL type. Next, the uncertainty associated with  $t_c$  in a realistic but nice case is quite large and is at the frontier of being exploitable. This strongly suggests that making predictions with real data is likely to yield worse uncertainties, since there is *a priori* no reason for the oscillations of real data to be systematically LPPL-based. Recent work that tries to estimate a probabilistic 20–80% confidence time window for  $t_c$  is certainly a step in the right direction. But since the time windows usually proposed are more optimistic than the reference case considered here, it is very likely that the method used underestimates the uncertainty on  $t_c$ . This is but an example of the problem of estimating parameter uncertainty from a single realisation of noise. As explained above,  $t_c$  may fluctuate very much when the time series given in input is changed slightly because it is a sloppy parameter; hence, the mere fact that it does fluctuate is not an indication *per se* that the variance of the fluctuations approximates correctly its

real uncertainty. Obtaining trustworthy predictions for sloppy parameters from a single time series is possible by Bayesian estimation (Brown and Sethna 2003, Frederiksen *et al.* 2004). Further work should look in this direction.

#### References

- Bastiaensen, K., Cauwels, P., Sornette, D., Woodard, R. and Zhou, W.-X., *The Chinese Equity Bubble: Ready to Burst*, 2009 (Scribd Inc.).
- Borodich, F.M., Renormalization schemes for earthquake prediction. *Geophys. J. Int.*, 1997, **131**(1), 171–178.
- Bothmer, H.C.G. and Meister, C., Predicting critical crashes? A new restriction for the free variables. *Physica A*, 2003, **320**, 539–547.
- Brée, D.S. and Joseph, N.L., Fitting the log periodic power law to financial crashes: A critical analysis. Cornell University, 2010.
- Brown, K. and Sethna, J., Statistical mechanical approaches to models with many poorly known parameters. *Phys. Rev. E*, 2003, **68**, 21904.
- Byrd, R., Schnabel, R. and Shultz, G., A trust region algorithm for nonlinearly constrained optimization. *SIAM J. Numer. Anal.*, 1987, **24**, 1152–1170.
- Chang, G. and Feigenbaum, J., A Bayesian analysis of log-periodic precursors to financial crashes. *Quant. Finance*, 2006, **6**, 15–36.
- Drożdż, S., Kwapien, J. and Oświecimka, P., Criticality characteristics of current oil price dynamics. *Acta Phys. Polon. A*, 2008a, **114**(4), 699–702.
- Drożdż, S., Kwapien, J., Oświecimka, P. and Speth, J., Current log-periodic view on future world market development. Cornell University, 2008b.
- Faillietaz, J., Funk, M. and Sornette, D., Icequakes as precursors of ice avalanches. Cornell University, 2009.
- Feigenbaum, J., A statistical analysis of log-periodic precursors to financial crashes. *Quant. Finance*, 2001, **1**, 346–360.
- Frederiksen, S., Jacobsen, K., Brown, K. and Sethna, J., Bayesian ensemble approach to error estimation of interatomic potentials. *Phys. Rev. Lett.*, 2004, **93**, 165501.
- Gazola, L., Fernandes, C., Pizzinga, A. and Riera, R., The log-periodic-AR(1)-GARCH(1,1) model of financial crashes. *Eur. Phys. J. B*, 2008, **61**, 355–362.
- Gutenkunst, R., Waterfall, J., Casey, F., Brown, K., Myers, C. and Sethna, J., Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput. Biol.*, 2007, **3**, 1871–1878.
- Huang, Y., Saleur, H. and Sornette, D., Reexamination of log periodicity observed in the seismic precursors of the 1989 Loma Prieta earthquake. *J. Geophys. Res.*, 2000, **105**(B12), 28111–28123.
- Jacobsen, E. How to predict crashes in financial markets with the log-periodic power law. Master’s Thesis, Mathematical Statistics, Stockholm University, 2009.
- Jiang, Z.-Q., Zhou, W.-X., Sornette, D., Woodard, R., Bastiaensen, K. and Cauwels, P., Bubble diagnosis and prediction of the 2005–2007 and 2008–2009 Chinese stock market bubbles. *J. Econ. Behav. Organiz.*, 2010, **74**, 149–162.
- Johansen, A., Ledoit, O. and Sornette, D., Crashes as critical points. *Int. J. Theor. Appl. Finance*, 2000, **3**, 219–225.
- Johansen, A. and Sornette, D., Bubbles and anti-bubbles in Latin-American, Asian and Western stock markets: An empirical study. *Int. J. Theor. Appl. Finance*, 2001, **4**, 853–920.
- Johansen, A. and Sornette, D., Shocks, crashes and bubbles in financial markets. *Brussels Econ. Rev. (Cahiers économiques de Bruxelles)*, 2010, **53**(2), 201–253 [Original title: Endogenous

- versus exogenous crashes in financial markets]. Available online at: <http://arxiv.org/abs/cond-mat/0210509> (accessed 10 February 2011).
- Kawada, Y., Nagahama, H., Yasuoka, Y., Omori, Y., Ishikawa, T., Tokonami, S. and Shinogi, M. Power-law and log-periodic changes in precursory phenomena prior to large earthquakes. In *The 33rd International Geological Congress*, 2008.
- Laloux, L., Potters, M. and Cont, R., Are financial crashes predictable? *Europhys. Lett.*, 1999, **45**, 1–5.
- Lillo, F. and Mantegna, R., Dynamics of a financial market index after a crash. *Physica A*, 2004, **338**, 125–134.
- Lin, L., Ren, R.E. and Sornette, D., A consistent model of ‘explorative’ financial bubbles with mean-reversing residuals. Cornell University, 2009.
- Malamud, B.D., Morein, G. and Turcotte, D.L., Log-periodic behavior in a forest-fire model. *Nonlinear Process. Geophys.*, 2005, **12**, 575–585.
- Matsushita, R., Gleria, I., Figueiredo, A. and Da Silva, S. On log-periodic crashes. Economics Working Paper Archive, Finance 0505007, University of Connecticut, 2005.
- Newman, W.I., Turcotte, D.L. and Gabrielov, A.M., Log-periodic behavior of a hierarchical failure model with applications to precursory seismic activation. *Phys. Rev. E*, 1995, **52**(5), 4827–4835.
- Nottale, L., Chaline, J. and Grou, P., On the fractal structure of evolutionary trees. In *Fractals in Biology and Medicine*, edited by G. Losa, D. Merlini, T. Nonnenmacher, and E. Weibel, Vol. 3, pp. 247–258, 2000 (Birkhäuser).
- Pralong, A., Birrer, C., Stahel, W.A. and Funk, M., On the predictability of ice avalanches. *Nonlinear Process. Geophys.*, 2005, **12**, 849–861.
- Rosenbrock, H.H., An automatic method for finding the greatest or least value of a function. *Comput. J.*, 1960, **3**, 175.
- Sornette, D., Discrete scale invariance and complex dimensions. *Phys. Rep.*, 1998, **297**(5), 239–270.
- Sornette, D. and Johansen, A., Large financial crashes. *Physica A*, 1997, **245**, 411–422.
- Sornette, D. and Johansen, A., Significance of log-periodic precursors to financial crashes. *Quant. Finance*, 2001, **1**, 452–471.
- Sornette, D., Johansen, A. and Bouchaud, J.-P., Stock market crashes, precursors and replicas. *J. Phys. I France*, 1996, **6**, 167–175.
- Sornette, D. and Sammis, C.G., Complex critical exponents from renormalization group theory of earthquakes: Implications for earthquake predictions. *J. Phys. I France*, 1995, **5**(5), 607–619.
- Sornette, D., Woodard, R. and Zhou, W.-X., The 2006–2008 oil bubble: Evidence of speculation, and prediction. *Physica A*, 2009, **388**, 1571–1576.
- Sornette, D. and Zhou, W.-X., Predictability of large future changes in major financial indices. *Int. J. Forecast.*, 2006, **22**, 153–168.
- Suchecki, K. and Holyst, J.A., Log-periodic oscillations in degree distributions of hierarchical scale-free networks. *Acta Phys. Polon. B*, 2005, **36**(8), 2499–2511.
- Tanaka, H., Log-periodic modulation to power law of change in the geoelectric potential difference observed on Niiijima Island prior to volcanic and seismic activation in the Izu Island region, Japan 2000. *Proc. Jap. Acad. Ser. B*, 2002, **78B**(9), 271–276.
- van Bothmer, H.G., Significance of log-periodic signatures in cumulative noise. *Quant. Finance*, 2003, **3**, 370–375.
- van den Bos, A. and Swarte, J.H., Resolvability of the parameters of multiexponentials and other sum models. *IEEE Trans. Signal Process.*, 1993, **41**, 313.
- Water, J.J., Casey, F.P., Gutenkunst, R.N., Brown, K.S., Myers, C.R., Brouwer, P.W., Elser, V. and Sethna, J.P., Sloppy-model universality class and the Vandermonde matrix. *Phys. Rev. Lett.*, 2006, **97**, 150601.
- Yana, W., Woodard, R. and Sornette, D. Diagnosis and prediction of tipping points in financial markets: Crashes and rebounds. Cornell University, 2010.
- Zhou, W. and Sornette, D., Is there a real-estate bubble in the US? *Physica A*, 2006, **361**(1), 297–308.
- Zhou, W.-X. and Sornette, D., Analysis of the real estate market in Las Vegas: Bubble, seasonal patterns, and prediction of the CSW indexes. *Physica A*, 2008, **387**, 243–260.